Methods of simplification for process of 3D animation production

Edin Pašović[1]
Supervised by: Jasminka Hasić [2]
Sarajevo School of Science and Technology
Bosnia and Herzegovina

Abstract

In the modern world, mass-produced computer graphics can be seen almost everywhere, ranging from TV commercials, movies, series, computer games, architecture, archeology, science, history, educational materials, medical research and countless other fields.

Digital animation studios like Pixar and Dreamworks employ thousands of artists and build large render farms and computer clusters for the purposes of rendering CGI (Computer Generated Graphics). Yet, there are thousands of small studios all over the world, producing all sorts of 3D graphics for small-production environments, and they mostly have up to 10 employees and no budget for rendering farms or other expensive computing solutions. They have to focus on using heurestics, shortcuts and alternative rendering methods to achieve results which, despite the low hardware resources for rendering, still keep the look of a high-quality production.

CR Categories: I.3.8 [Computer Graphics]: Applications
Keywords: visual perception, selective attention, GPGPU, parallel processing

1 Visual Language

The way 3D artists create and visualize the scenes themselves can have a great impact on the rendering speed and quality of animations.

A visual language is a set of practices by which images (such as diagrams, professional photos or even motion 3D movies) can be used to communicate concepts and stories. Whether it's a diagram, a professional photograph or a map, they all represent the use of visual language. The structural units of visual language are: shape, line, texture, pattern, color, motion, direction, orientation, scale, space, angle and proportion.[3]

Most known texts on subject of visual language in movies and motion pictures were written by Rudolf Arnheim, a German-born author, art and film theorist and perceptual psychologist. In my project "Greyworld", several of those techniques were used to demonstrate examples of how an efficient usage of visual language can significantly reduce workflow and rendering times.

2 Efficiency in visual language

Recent studies in human perception [Koch, C. and Ullman, 1985; Niebur, E. and Koch, C. 1996; Cater, K., Chalmers, A. and Ward, G. 2003] show increasing success in prediction of eye movements and focus of subjects watching motion pictures and give us further insights into the concepts Rudolf Arnheim first mentioned in his book "Art and Visual Perception: A Psychology of the Creative Eye" in 1954 and later explored in "Visual Thinking" in 1969 and "New Essays on the Psychology of Art" in 1986.

Among many insights into the way the common viewer percieves motion pictures, there is a number of those which, when used in right places, could help us decrease scene complexity or time invested into the creation of the scene in a 3D animation.

Next several examples are used in everyday work in modern animation workflows, and they were not quoted from any particular book, but were acquired by the author of this paper as a result of insight after several years of working in the graphic design and animation industry and listening to the experiences of seniors in that industry. Most of them are also explained in some of the professional DVD tutorial courses [HCWVEFD 2009; HCWMCHEBS 2004].

2.1 Foreground and Background

A large percentage of actual action in motion pictures happens in foreground, thus drawing the focus of the viewer much more on the foreground than on the background.

If there is nothing of importance happening in the background during the scene, the viewer will be

[1] edinpasovic@gmail.com

[2] jh@ssst.edu.ba

gradually less and less focused on the background. This should basically allow us to render the background details with far less detail once we're several seconds into the scene, because by then the viewer will lose interest in analysing the background (like he did for the first few seconds), and he will shift it towards the action in the foreground, not noticing the decline in rendering quality in the background, as long as nothing drastically changes. [HCWMCHEBS 2004]

2.2 Depth of Field

The Depth of Field (DOF) effect, when used inside the 3D software in scenes with high complexity, can increase rendering process up to five times[4].

There is a new method of visualisation, which consists of dividing the scene into several layers, depending on the distance of objects, rendering them without DOF effect, and bringing them together later in compositing software, where DOF simulation will be applied on layers instead of actual objects. While the results are very similar in appearance, the rendering times are significantly decreased[3]. Today this method is mostly being used in the latest generation of game-consoles, such as X-Box 360 and Playstation 3, to simulate DOF in games while maintaining constant frame-rate. [HCWVEFD 2009]

2.3 Moving and static objects

Moving objects tend to draw more attention on themselves than the static objects. Objects that move towards us tend to draw more attention on themselves than the objects that move away from us.

In anthropology this is explained as an old human instinct still left from times when humans used to hunt animals in oder to survive. This behavior can be exploited in order to decide which objects should be modeled and rendered more carefully and which ones require less polygons and simpler rendering procedures, based on their movement. [Glencross M., Chalmers A.G., Lin M.C., Otaduy M.A., Gutierrez D. 2006]

2.4 Color

Human eye is capable of color vision thanks to photosensitive Retinal Ganglion Cells. However, these cells don't percieve all the colors with equal sensitivity.[5] We are most sensitive to red color, less sensitive to green color and even less sensitive to blue color. This is why red and yellow objects (yellow color in RGB color model is combination of red and green) in the scene will draw more attention than the others.

This information can be used to classify which parts of the scene, based on their color, should be given more or less importance in modeling and rendering. [R. Arnheim 1974]

2.5 Composition

Most people in Europe, North and South America, Australia and northern parts of Asia use alphabets written from left to right. This affects their perception so that they observe and analyze static and motion pictures in similar manner - from left to right.

This can be seen in many traditional movie scenes, where camera is following one or more characters from the side during the walking scene - they mostly tend to go from left to right. Video games follow the same analogy, starting from the first "Super Mario" side-scrolling game to the most modern side-scrolling game - characters start at point A on the left side and have to fight their way to the point B which is in most cases located on the right side.

A common viewer tends to focus on what's in front of the character and where the character is going - so in those cases he is mostly focusing on the right side, while the parts of the scene which stay behind recieve much less focus once they are behind characters back. This, again, can be used to gradually reduce the rendering quality of those parts of the scene which are moving towards the left side of the screen. [HCWMCHEBS 2004]

3 Case Study: "Greyworld"

The short 3D animation "Greyworld" was built using both layered 2D and 3D techniques of animation. Since this was a project one single animator was working on, in a limited amount of time (three-and-a-half months), using only one non-workstation computer (Apple iMac, Duo Core 2,6 Ghz, 2 GB RAM, ATI 2660 HD GPU), the goal was to produce an animation while focusing only on important visual details, and invest time and work on areas in regard of how visible and how important they will be to the final audience.

The whole final animation consists of 46 shots distributed through 10 scenes, with a duration of 5 minutes and 40 seconds. During the various production stages, six candidates, with no background or actual involvement in animation or production, were shown segments of the animation and asked if they see any mistakes or details that need additional refining and notices were taken about what they focused on and what they felt was important in the animation.

[4] Actual number depends on the complexity of the scene

5 human eye - ganglion cells." Encyclopædia Britannica. 2010. Encyclopædia Britannica Online. 02 Mar. 2010 <http://www.britannica.com/EBchecked/topic/199272/eye

4 Experiment methodology

Members of the audience were shown single scenes from the animation during the production stage and asked what they find important in the scene and if there are any corrections needed. I always showed them the more complex animation first and the simplified version second.

Questionnaire had the following form:

SCENE A1
Which scene seemed more complex
a) first,
b) second
c) both were equally complex

SCENE A2
Did you notice any difference in fog between the two scenes?
a)yes, clearly
b) yes, but nothing significant
c) not at all

SCENE B
Which model seemed more complex
a) first one
b) second one
c) both were equally complex

SCENE C
Which model seemed more complex
a) first one
b) second one
c) both were equally complex

Answer	a)	b)	c)
SCENE A1	2 (33%)	0	4 (66%)
SCENE A2	0	3 (50%)	3 (50%)
SCENE B	2 (33%)	0	4 (66%)
SCENE C	0	0	6 (100%)

Studying their answers, I realized that they mostly focus on the foreground, paying almost no attention to the background.

They noticed the difference between scenes where I eliminated the background and the scenes where I added it, but they were so focused on the main characters and their actions that they didn't notice the difference between scenes with detailed background and the same scenes with the same background in much lower detail.

Therefore we decided to put a 360 image sphere of an industrial town as a background and skip all the time and work which would be spent into making 3D models of background buildings, distributing and rendering them, effectively eliminating several days of work. In order to make this simplification of the background less noticeable, a depth-of-field effect was used to add a slight blur to the background.

Figure 1: Example frames from the animation "Greyworld"

There are two main characters, the boy and the robot. Since the boy bears much more similarity with human beings than the robot, audience members were noticing even the smallest mistakes and asked for more details on the model of the boy, when compared to the robot. This is due to the natural human instinct to pay much more attention to detail when watching a human-like being than when observing an animal, robot or a shape. Therefore, the final version of the boy was modeled in 11.000 polygons while the robot didn't take more than 700 polygons.

Also regarding the physical features of the boy's model, viewers paid much more attention to the head and the face than to the other parts and limbs. So in the final model, more than 4.000 of 11.000 polygons were used just to model the face region.

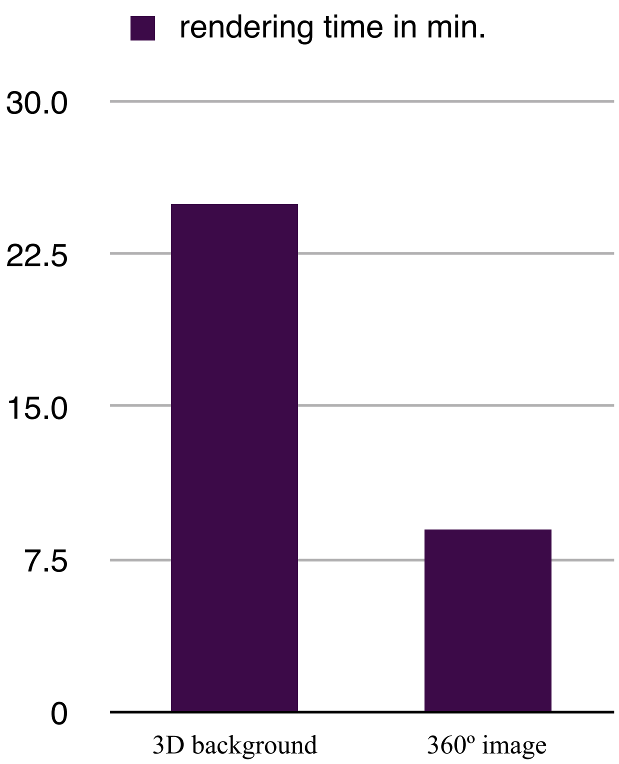
Weather effects and particle effects, such as fog and smoke coming from the chimneys were noticed by the viewers, however they couldn't see any significant difference between 3D volumetric fog and 2D fog made using noise maps in After Effects (a significantly faster method) except in cases where we told them to pay more attention to the fog. The same results came up in the case of smoke from the chimneys in the intro scene. Without being previously warned about the difference between two types of smoke, observers wouldn't notice the difference between the scene where the smoke was simulated through volumetric 3D particles, and the scene where it was just applied as a simple 2D particle effect in post-production stage.

5 Results

Simplification of less important elements of animation led to a noticeable decrease in rendering times.

5.1 Scene A1.

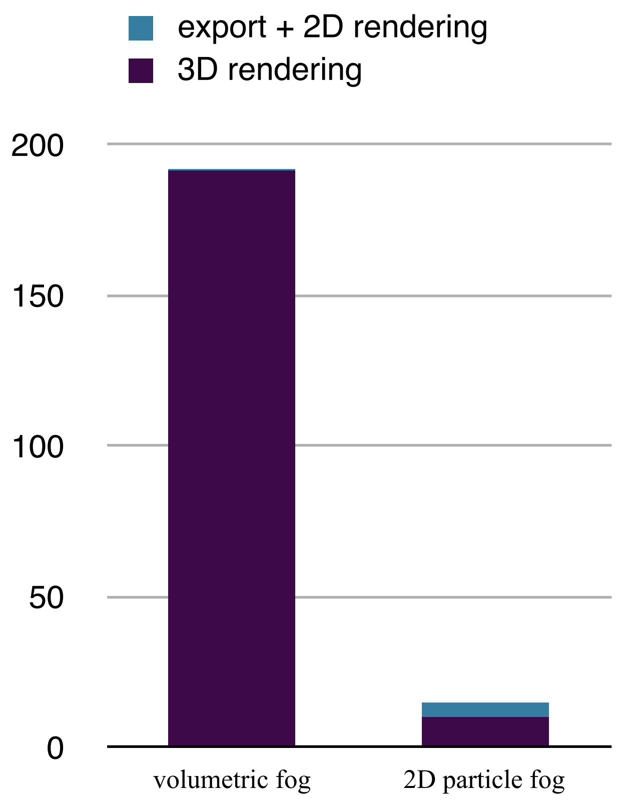
An example scene with 360º image in the background was rendered in 9 minutes, while the same scene with 3D modeled buildings in the background took 25 minutes to render. (only 36% of the original time required). Four of six candidates (66%) noticed no difference before they were told about it.



5.2 Scene A2

In the same example scene, we used volumetric fog, generated in Cinema 4D software, to enhance the mood, and it took over 190 minutes to render, but when we rendered the scene without fog (9 minutes) and exported it to After Effects (in FBX format with 3D cameras and lights), where we added 2D fog with the particle simulation system, it just took another 5,5 minutes for it to render, so the final processing time was 14,5 minutes, compared to 190 minutes with volumetric fog. (7,63% of the original time required).

Three of the candidates (50%) didn't notice the change in the quality of fog before they were told about it, while another three (50%) noticed it but labeled it as not significant in magnitude. But even after noticing the difference in two fog rendering methods, their focus drifted off the fog and to the other elements of the scene. 30 seconds after the fog was introduced into the scene, it was still there and all six of the participants (100%) could confirm its presence, but they were not noticing any details or quality changes regarding the fog anymore, it was just background information now. This led me to conclude that fog could be rendered in full quality and volumetric for the first several seconds of its introduction as an element of the scene, and then lowered in quality and complexity in order to decrease rendering times.



5.3 Scene A3

In this example scene boy is walking across the lawn and then comes to a halt. Two polygonal meshes for the model of the boy were used in two different version of the scene. The complex one consists of 34.000 polygons, the simpler one counts less than 11000 polygons, optimized so that the most complexity is in the face of the boy.

Six of participants (100%) didn't notice any difference between two versions of the scene while the boy was walking (being in motion). However, two of them (33%) noticed this difference after he came to a halt (stopped moving), but only during the second viewing, after they have been told of the difference. This led me to conclude that simpler versions of the models can be used in moving sequences, while the more detailed (and processor-hungry) models may be left only for still scenes if required.

SCENE C

The same example scene was used as in "SCENE B", however this time I used the 9000 polygon and the 700 polygon model of the robot.

Six of the participants (100%) didn't notice any difference between models in the scenes, and even after being told about it, and after watching still frames from the scene, they couldn't find any considerable difference. This led me to conclude that people pay less attention to detail when watching a non-human creature or robot, than when watching a human or human-like creature.

6 Conclusions and feature work

Using methods such as visual language or experiments with a live audience helps to decide which elements of the scene require a specific level of quality, effectively reducing work and processing times for a significant amount. Further insight in this field could result in classifications, such as a detailed table of specific scene elements showing their relative level of importance and thus saving time and money when working on small and medium-sized 3D animation projects, especially if the animators are less experienced.

While this approach might seem very subjective, with no clear formulas or algorithms, there are still visible patterns in perception of most viewers, which can be statistically measured and compared. Further research could give us answers on how much we have to go into detail with other animation tasks that are heavy on processing requirements, such as fluid animation, cloth dynamics, hair dynamics and other particle dynamics, when working on small-to-medium budget projects, resulting in a clearly defined table of priorities.

Acknowledgements

We would like to thank Andrei Ferko for his professional support during the production stage, Muhamed Kafedzic and Matej Srepfler for sharing their experience on compositing and special effects simplification, and Dino Karadza, Albin Brkic, Dijana Brkic, Elma Selman, Matej Srepfler and Mirza Coric for reviews during the animation production stage.

References

[1] Isaac V. Kerlow. *The Art of 3D Computer Animation and Effects (Third Edition).* Wiley Publishing Inc. 2005.

[2] Rudolf Arnheim. *Art and Visual Perception: A Psychology of the Creative Eye (Second Edition).* University of California Press, 1974.

[3] Glencross M., Chalmers A.G., Lin M.C., Otaduy M.A., Gutierrez D. *Exploiting Perception in High-Fidelity Virtual Environments.* ACM SIGGRAPH, July 2006.

[4] Andrea Di Blas, Tim Kaldewey. *Data Monster - Why graphics processors will transform database processing.* Article on ieee.spectrum website, September 2009.

[5] Rolf Herken. *The Future of Rendering.* NVISION08 presentation, 2008

[6] Rui Wang, Kun Zhou, Minghao Pan, Hu-Jun Bao. *An Efficient GPU-based Approach for Interactive Global Illumination.* Siggraph presentation, 2009

[7] Koch, C. and Ullman, S. *Shifts in selective visual attention: towards the underlying neural circuitry.* Human Neurobiology, 1985.

[8] Niebur, E. and Koch, C. *Control of Selective Visual Attention: Modeling the `Where' Pathway.* Neural Information Processing Systems, 1996.

[9] Cater, K., Chalmers, A. and Ward, G. *Detail to attention: Exploiting visual tasks for selective rendering*. Proceedings of the Eurographics Symposium on Rendering, 2003.

[10] Hasic, J., and Chalmers, A. *Visual attention for influencing the perception of virtual environment.* Spring Conference on Computer Graphics, ACM SIGGRAPH, April 2006.

[11] Hasic, J., and Chalmers, A. *Visual attention for significantly influencing the perception of virtual environments.* Winter School of Computer Graphics WSCG2007 Proceedings, 2007.

[12] NVIDIA Corporation. *Fermi Compute Architecture Whitepaper.* Article on www.nvidia.com

[13] HCWVEFD 2009 - Hollywood Camera Work. *Visual effects for Directors* DVD edition, 2009

[14] HCWMCHEBS 2004 - Hollywood Camera Work. *The Master Course in High-End Blocking & Staging.* DVD edition, 2009